Module Thematic Analysis using Comment and ChatGPT
By Dr Nurul Ashikin Izhar

This research analysis metho is a part of Artificial Intelligence (AI)- Powered Research Assistant for Beginner.

1.0 Conducting analysis using Comments and ChatGPT

This is an alternative idea on conducting analysis in using available and free software. Comments here refer to the Comment function available in Microsoft Word and ChatGPT referred to the generative artificial intelligence (AI).

The example used here is data that been collected using Google Form.

1.1 Steps to conduct the analysis

Step 1: Data Extraction and Initial Coding

The first step in thematic analysis involves extracting the qualitative data and applying initial codes. For instance, if the data is collected through an online platform such as Google Forms, the responses should be exported in a CSV format. This raw data is then transferred to a more workable format, such as Microsoft Word, where researchers can easily engage with it.

Once the data is organized, researchers must familiarize themselves with the content by reading through the responses. This immersion allows them to begin the initial coding process. Using the comment function in Word, researchers assign codes by highlighting recurring words, phrases, or ideas that stand out in the text. These codes represent key elements or concepts expressed by the participants. An essential part of this process is collaboration—typically, at least three researchers work independently to apply open coding to the data. This collaborative approach not only enhances the richness of the analysis but also increases inter-rater reliability, which ensures that different researchers consistently apply the same codes across the dataset.

Summary step 1:
Data Extraction and Initial Coding

- Export Responses: Begin by exporting the responses from Google Forms into a CSV file.
- Transfer Data: Import the CSV data into Microsoft Word to organize the text for analysis.
- Initial Review: Researchers should read through the responses to familiarize themselves with the data.
- Assign Codes: Use the Comment function in Word to assign initial codes to segments of the text. This is done by highlighting recurring words, ideas, or concepts in the data.
- Collaborative Coding: At least three researchers should collaboratively engage in open coding, where they apply codes independently.
- Recheck Process: Conduct a review of the coding process to ensure inter-rater reliability, meaning that the researchers agree on the application of the codes.

Step 2: Code Extraction

Once the initial coding is completed, the next task is to extract all the codes for further analysis. Microsoft Word's Macro function can be used to automatically gather all the comments or codes that have been applied during the coding process. This method saves time and organizes the data more efficiently, allowing the researchers to focus on theme development without losing any critical codes.

To enable the macros function, specific coding needs to be set up, which can be found in **Appendix A**. In Microsoft Word, follow these steps:

1. Go to **View** > **View Macros** > **Create**.
2. Copy and paste the coding from **Appendix A** into the macro editor.
3. Click **Save** and then close the editor window.

To run the macro and extract the data, go to **Macros** > **View Macros** > **Run**. The data will be extracted into a new document.

Summary Step 2:
Extract Comments:

- Use a Macro function in Word to extract all the codes (comments) that were assigned during the initial coding phase. This will help in organizing the codes for further analysis.

Step 3: AI-Assisted Theme Generation

At this stage, researchers move from manual coding to leveraging AI for theme generation. The coded data extracted in the previous step is uploaded to an AI tool like ChatGPT for assistance. AI plays a crucial role in categorizing the codes into subthemes, which represent related ideas or concepts. This process groups similar codes together, which helps in narrowing down the broad dataset into more focused and manageable themes.

Once the subthemes are generated, the AI can synthesize them into overarching themes that encapsulate the broader insights emerging from the data. However, it's important to note that while AI can speed up the process, the human researcher must review the AI-generated themes to ensure they are meaningful and relevant to the research questions and context. This review is critical because AI, while helpful, may not fully capture the nuances of the data without proper oversight.

Summary Step 3:
AI-assisted Theme Generation

- Upload Data to ChatGPT: Once the coded data is extracted, upload it to ChatGPT for assistance with theme generation.
- Categorize Codes: Prompt ChatGPT to categorize the codes into subthemes, grouping related codes together based on patterns or similarities.
- Synthesize Themes: ChatGPT will help synthesize these subthemes into overarching themes that represent broader concepts in the data.
- Review by Researchers: The AI-generated subthemes should then be reviewed by researchers to ensure they align with the research objectives and are meaningful within the context of the data.

Step 4: Iterative Analysis and Refinement

The fourth step in the thematic analysis process is focused on refining the themes that were generated with the help of AI. At this point, researchers must ensure the relevance and precision of the subthemes by evaluating how well they represent the underlying data. This iterative process involves a detailed review of the themes to check for any inaccuracies or inconsistencies.

Researchers can prompt ChatGPT to consolidate similar or overlapping subthemes into broader, more comprehensive themes. This step helps in eliminating redundancy and ensures that each theme captures a distinct aspect of the data. Additionally, further adjustments can be made as needed, allowing the researchers to refine their analysis to a point where the themes are not only accurate but also actionable and insightful.

Summary of Step 4:
Iterative Analysis and Refinement

- Ensure Relevance and Precision: Researchers should ensure the relevance and precision of the subthemes by checking how well they represent the data.
- Consolidate Themes: Prompt ChatGPT to consolidate subthemes into broader, more comprehensive themes, eliminating redundancy or overlapping concepts.
- Adjustments: Implement further refinements or adjustments to the themes based on researcher input and the evolving understanding of the data.

Final Step: Identifying Final Themes

The final stage of the thematic analysis involves the identification of the final themes. After refining and consolidating the subthemes, the final set of themes represents the key findings from the qualitative data. These themes are the result of a meticulous process that combines human judgment with AI-assisted efficiency. Once finalized, the themes can be presented in a research report, providing valuable insights into the data and addressing the research questions effectively.

Figure 1 illustrates the steps in flowchart manners

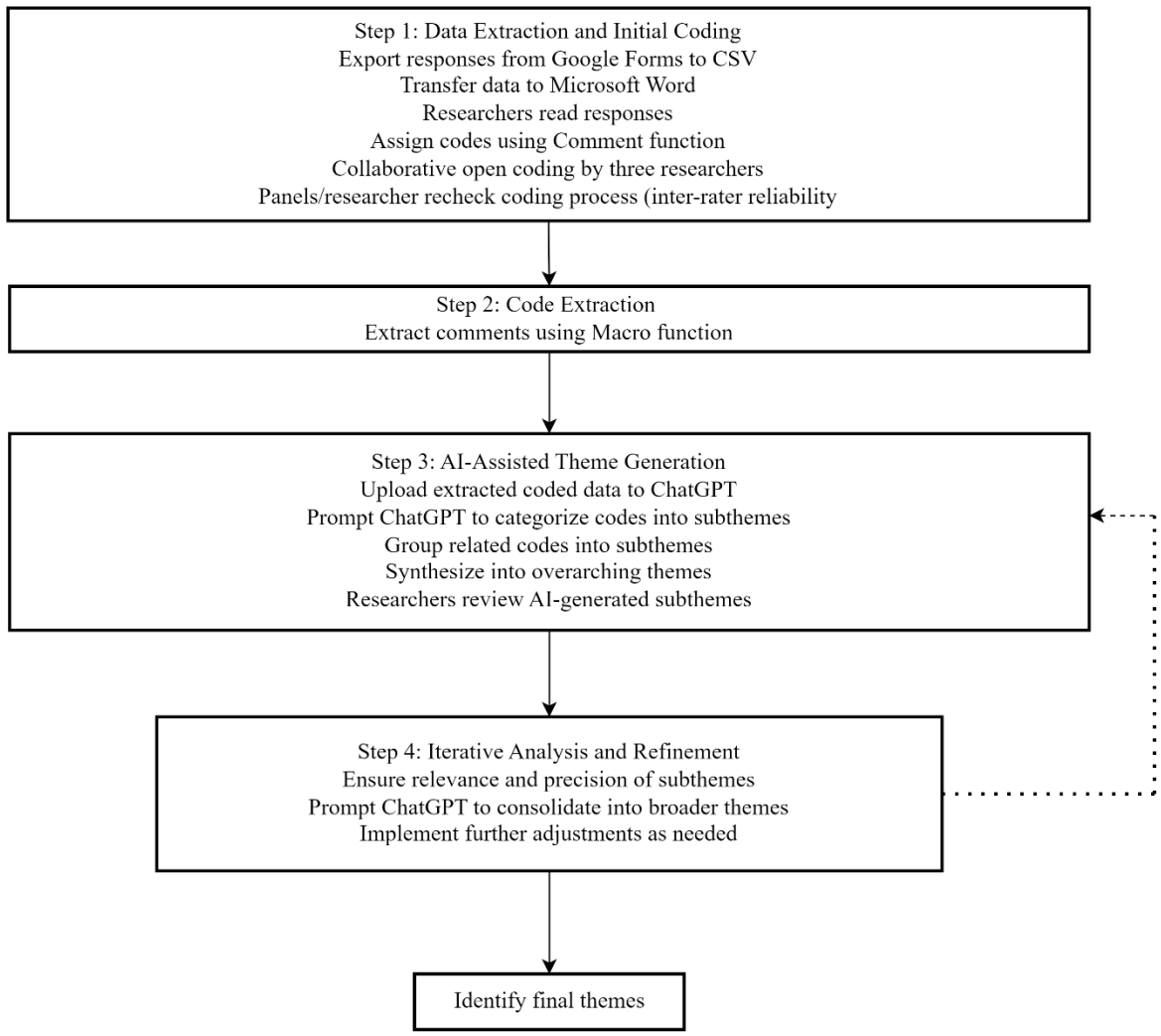

Figure 1 Thematic analysis using Comment and ChatGPT

This method presents a valuable alternative for conducting thematic analysis, offering a blend of traditional qualitative analysis techniques with the efficiency of modern AI tools. By integrating AI tools like ChatGPT into the thematic analysis process, researchers can streamline data analysis while still maintaining depth and rigor. This alternative approach ensures that key themes are identified, refined, and organized to provide meaningful insights. Additionally, the use of AI accelerates the coding process and aids in uncovering patterns and relationships that may be missed through manual analysis alone. This method allows researchers to conduct a comprehensive and reliable thematic analysis, contributing to a deeper understanding of the data and enabling actionable conclusions.

1.2 Suggestion Prompt for Qualitative analysis using generative artificial intelligence (such as ChatGPT)

a. For Initial Coding:

1. Here is a transcript of interviews/survey responses. Can you identify recurring words, phrases, or ideas that stand out?
2. Please generate a list of codes based on this qualitative data. Highlight key ideas or categories mentioned by participants.
3. Based on this text, what are the most frequent topics or issues raised by respondents?

b. For Pattern Identification and Theme Development:

4. Please group similar codes or ideas together and suggest overarching themes from this data.
5. Identify relationships between the different codes and propose broader categories or themes emerging from the dataset.
6. Analyze this data for common patterns or trends. What recurring themes or insights do you observe?

c. For Axial Coding:

7. How are these codes related? Can you create sub-themes or categories based on these relationships?
8. Please help me conduct axial coding by identifying connections between the themes and organizing them into broader categories.
9. Look at the following themes and suggest how they are connected. Can you refine them into sub-categories?

d. For Theme Refinement:

10. Based on these identified themes, can you help refine them into more focused or actionable categories?
11. Can you simplify or collapse these themes into fewer but more comprehensive ones?
12. Evaluate the consistency of these themes. Are there any overlaps or redundancies that can be merged?

e. For Data Summary:

13. Summarize the key findings from this qualitative data and explain the major themes or categories that emerged.
14. Create a summary of the themes, including supporting quotes or examples from the data for each theme.
15. Based on this data, what insights or conclusions can be drawn about the participants' attitudes or experiences?

2.0 Conclusion

Thematic analysis is a powerful tool for analyzing qualitative data, allowing researchers to systematically identify patterns and themes within complex narratives. This module has provided a comprehensive guide to conducting thematic analysis, combining traditional methods with modern AI assistance through tools like ChatGPT. By leveraging this approach, researchers can enhance the efficiency of their analysis while maintaining the rigor and depth required for meaningful insights.

The hybrid approach outlined in this steps—combining manual and AI-driven analysis—offers flexibility and innovation in qualitative research. By utilizing AI tools, researchers can reduce time spent on labour-intensive tasks like coding, while still maintaining control over the interpretive process. Importantly, this approach opens the door to more extensive and nuanced analysis, enabling the discovery of insights that might otherwise remain hidden in large datasets.

References, *Rujukan*

Abdekhodaie, E., Hatami, J., Bahrami Ehsan, H., & Kormi-Nouri, R. (2018). WordCommentsAnalyzer: A windows software tool for qualitative research. *F1000Research*, *7*, 1–17. https://doi.org/10.12688/f1000research.14819.1

Chew, R., Bollenbacher, J., Wenger, M., Speer, J., & Kim, A. (2023). LLM-Assisted Content Analysis: Using Large Language Models to Support Deductive Coding. *ArXiv*. https://doi.org/https://doi.org/10.48550/arXiv.2306.14924.

Şen, M., Şen, Ş. N., & Şahin, T. G. (2023). A New Era for Data Analysis in Qualitative Research: ChatGPT! *Shanlax International Journal of Education*, *11*(S1-Oct), 1–15. https://doi.org/10.34293/education.v11is1-oct.6683

Appendix A Coding for Macros

```
Sub ExportCommentsToNewDocument()
    ' This macro extracts comments from the active document into a new document.
    ' A table is created with columns for page, line, comment text, author, and date.

    Dim sourceDoc As Document
    Dim targetDoc As Document
    Dim commentTable As Table
    Dim commentCount As Long
    Dim i As Long
    Dim headerTitle As String

    headerTitle = "Comments Extraction Report"
    Set sourceDoc = ActiveDocument
    commentCount = sourceDoc.Comments.Count

    ' Check if the document has comments
    If commentCount = 0 Then
        MsgBox "No comments found in the active document.", vbOKOnly, headerTitle
        Exit Sub
    End If

    ' Confirm extraction
    If MsgBox("Do you want to export all comments to a new document?", _
            vbYesNo + vbQuestion, headerTitle) <> vbYes Then
        Exit Sub
    End If

    Application.ScreenUpdating = False
```

```
    ' Create new document
    Set targetDoc = Documents.Add
    targetDoc.PageSetup.Orientation = wdOrientLandscape

    ' Insert a table to organize comments
    Set commentTable = targetDoc.Tables.Add(Range:=targetDoc.Range, _
                NumRows:=commentCount + 1, NumColumns:=6)

    ' Add header information to the new document
    With targetDoc.Sections(1).Headers(wdHeaderFooterPrimary).Range
        .Text = "Source Document: " & sourceDoc.FullName & vbCrLf & _
            "User: " & Application.UserName & vbCrLf & _
            "Date: " & Format(Date, "MMMM d, yyyy")
    End With

    ' Customize the styles for better readability
    With targetDoc.Styles(wdStyleNormal)
        .Font.Name = "Calibri"
        .Font.Size = 11
    End With

    ' Set up the table headings
    With commentTable.Rows(1)
        .Range.Font.Bold = True
        .Cells(1).Range.Text = "Page"
        .Cells(2).Range.Text = "Line"
        .Cells(3).Range.Text = "Commented Text"
        .Cells(4).Range.Text = "Comment"
        .Cells(5).Range.Text = "Author"
        .Cells(6).Range.Text = "Date"
    End With

    ' Extract comments and populate the table
    For i = 1 To commentCount
```

```
        With commentTable.Rows(i + 1)
            .Cells(1).Range.Text = sourceDoc.Comments(i).Scope.Information(wdActiveEndPageNumber)
            .Cells(2).Range.Text = sourceDoc.Comments(i).Scope.Information(wdFirstCharacterLineNumber)
            .Cells(3).Range.Text = sourceDoc.Comments(i).Scope.Text
            .Cells(4).Range.Text = sourceDoc.Comments(i).Range.Text
            .Cells(5).Range.Text = sourceDoc.Comments(i).Author
            .Cells(6).Range.Text = Format(sourceDoc.Comments(i).Date, "dd-MMM-yyyy")
        End With
    Next i

    Application.ScreenUpdating = True

    ' Display a message with the count of comments extracted
    MsgBox commentCount & " comments have been extracted.", vbOKOnly, headerTitle

    ' Clean up objects
    Set sourceDoc = Nothing
    Set targetDoc = Nothing
    Set commentTable = Nothing
End Sub
```